

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
17 June 2004 (17.06.2004)

PCT

(10) International Publication Number  
**WO 2004/051627 A1**

(51) International Patent Classification<sup>7</sup>: **G10L 19/08**

(21) International Application Number:  
PCT/IB2003/005019

(22) International Filing Date:  
6 November 2003 (06.11.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
02080002.5 29 November 2002 (29.11.2002) EP

(71) Applicants (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL). **SLUIJTER, Robert, J.** [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **DEN BRINKER, Albertus, C.** [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **GERRITS, Andreas, J.** [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(74) Agent: **GROENENDAAL, Antonius, W., M.**; Philips Intellectual Property & Standards, Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR,

KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

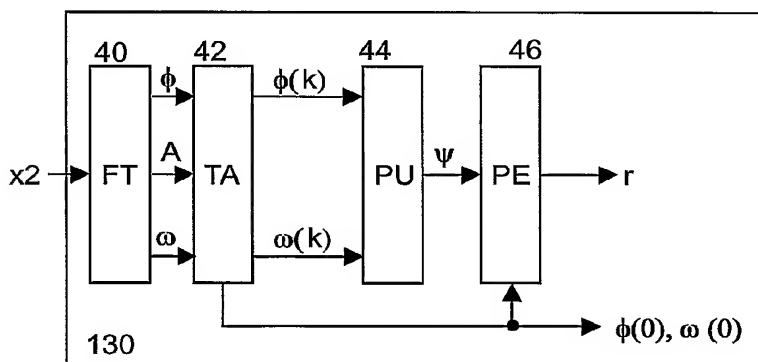
— as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

**Published:**

— with international search report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: AUDIO CODING



(57) Abstract: Coding of an audio signal represented by a respective set of sampled signal values for each of a plurality of sequential segments is disclosed. The sampled signal values are analysed (40) to determine one or more sinusoidal components for each of the plurality of sequential segments. The sinusoidal components are linked (42) across a plurality of sequential segments to provide sinusoidal tracks. For each sinusoidal track, a phase comprising a generally monotonically changing value is determined and an encoded audio stream including sinusoidal codes (r) representing said phase is generated (46).

WO 2004/051627 A1

## Audio coding

## FIELD OF THE INVENTION

The present invention relates to coding and decoding audio signals.

## BACKGROUND OF THE INVENTION

5 Referring now to Figure 1, a parametric coding scheme in particular a sinusoidal coder is described in PCT Patent Application No. WO01/69593. In this coder, an input audio signal  $x(t)$  is split into several (overlapping) segments or frames, typically of length 20ms. Each segment is decomposed into transient, sinusoidal and noise components. (It is also possible to derive other components of the input audio signal such as harmonic  
10 complexes although these are not relevant for the purposes of the present invention.)

In the sinusoidal analyser 130, the signal  $x_2$  for each segment is modelled using a number of sinusoids represented by amplitude, frequency and phase parameters. This information is usually extracted for an analysis interval by performing a Fourier Transform (FT) which provides a spectral representation of the interval including: frequencies;  
15 amplitudes for each frequency; and phases for each frequency where each phase is in the range  $\{-\pi, \pi\}$ . Once the sinusoidal information for a segment is estimated, a tracking algorithm is initiated. This algorithm uses a cost function to link sinusoids with each other on a segment-to-segment basis to obtain so-called tracks. The tracking algorithm thus results in sinusoidal codes  $C_s$  comprising sinusoidal tracks that start at a specific time instance, evolve  
20 for a certain amount of time over a plurality of time segments and then stop.

In such sinusoidal coding, frequency information is usually transmitted for the tracks formed in the encoder. This can be done cheaply, since tracks are defined as having a slowly varying frequency and, therefore, frequency can be transmitted efficiently by time-differential encoding. (In general, amplitude can also be encoded differentially over time.)

25 In contrast to frequency, phase transmission is viewed as expensive. In principle, if the frequency is (nearly) constant, phase as a function of the track segment index should adhere to a (nearly) linear behaviour. However, when it is transmitted, phase is limited to the range  $\{-\pi, \pi\}$  as provided by the Fourier Transform. Because of this modulo  $2\pi$

representation of phase, the structural inter-frame relation of the phase is lost and, at first sight appears to be a white stochastic variable.

However, since the phase is the integral of the frequency, the phase need, in principle, not be transmitted. This is called phase continuation and reduces the bit rate significantly.

In phase continuation, only the frequency is transmitted and the phase is recovered at the decoder from the frequency data by exploiting the integral relation between phase and frequency. It is known, however, that the phase can only be approximately recovered using phase continuation. If frequency errors occur, due to measurement errors in the frequency or due to quantisation noise, the phase, being reconstructed using the integral relation, will typically show an error having the character of a drift. This is because frequency errors have an approximately white noise character. Integration amplifies low-frequency errors and, consequently, the recovered phase will tend to drift away from the actually measured phase. This leads to audible artifacts.

This is illustrated in Figure 2(a) where  $\psi$  and  $\Omega$  are the real frequency and phase for a track. In both the encoder and decoder frequency and phase have an integral relationship represented by I. The quantisation process in the encoder is modelled as an additive white noise  $n$ . In the decoder, the recovered phase  $\hat{\psi}$  thus includes two components: the real phase  $\psi$  and a noise component  $\varepsilon_2$ , where both the spectrum of the recovered phase and the power spectral density function of the noise  $\varepsilon_2$  have a pronounced low-frequency character.

Thus, it can be seen that in phase continuation, since the recovered phase is the integral of a low-frequency signal, the recovered phase is a low-frequency signal itself. However, the noise introduced in the reconstruction process is also dominant in this low-frequency range. It is therefore difficult to separate these sources with a view to filtering the noise  $n$  introduced during encoding.

The present invention attempts to mitigate this problem.

## DISCLOSURE OF THE INVENTION

According to the present invention there is provided a method according to claim 1.

According to the invention the prior art sinusoidal coding technique is reversed i.e. phase rather than frequency is transmitted. In the decoder, the frequency can be approximately recovered from the quantised phase information using finite differences as an

approximation for differentiation. The noise component of the recovered frequency has a pronounced high-frequency behaviour under the assumption that the noise introduced by the phase quantisation is nearly spectrally flat. This is illustrated in Figure 2(b), where within the encoder and the decoder, frequency is represented as the differential (D) of phase. Again,  
5 noise  $n$  is introduced in the encoder and so in the decoder, the recovered frequency  $\hat{\Omega}$  includes two components: the real frequency  $\Omega$  and a noise component  $\varepsilon_4$ , where the frequency is nearly a DC signal and the noise is mainly in high-frequency range. However, since the underlying frequency has a low-frequency behaviour and the added noise a high-frequency behaviour, the noise component  $\varepsilon_4$  of the recovered frequency can be reduced by  
10 low-pass filtering.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows an audio coder in which an embodiment of the invention is implemented;

15 Figures 2(a) and 2(b) illustrate the relationship between phase and frequency in prior art systems and in audio systems according to the present invention respectively;

Figures 3(a) and 3(b) show a preferred embodiment of a sinusoidal coder component of the audio coder of Figure 1;

Figure 4 shows an audio player in which an embodiment of the invention is  
20 implemented; and

Figures 5(a) and 5(b) show a preferred embodiment of a sinusoidal synthesizer component of the audio player of Figure 4; and

Figure 6 shows a system comprising an audio coder and an audio player according to the invention.

#### DESCRIPTION OF THE PREFERRED EMBODIMENT

Preferred embodiments of the invention will now be described with reference to the accompanying drawings wherein like components have been accorded like reference numerals and, unless otherwise stated perform a like function. In a preferred embodiment of  
30 the present invention, the encoder 1 is a sinusoidal coder of the type described in PCT Patent Application No. WO 01/69593, Figure 1. The operation of this prior art coder and its corresponding decoder has been well described and description is only provided here where relevant to the present invention.

In both the prior art and the preferred embodiment, the audio coder 1 samples an input audio signal at a certain sampling frequency resulting in a digital representation  $x(t)$  of the audio signal. The coder 1 then separates the sampled input signal into three components: transient signal components, sustained deterministic components, and sustained stochastic components. The audio coder 1 comprises a transient coder 11, a sinusoidal coder 13 and a noise coder 14.

The transient coder 11 comprises a transient detector (TD) 110, a transient analyzer (TA) 111 and a transient synthesizer (TS) 112. First, the signal  $x(t)$  enters the transient detector 110. This detector 110 estimates if there is a transient signal component and its position. This information is fed to the transient analyzer 111. If the position of a transient signal component is determined, the transient analyzer 111 tries to extract (the main part of) the transient signal component. It matches a shape function to a signal segment preferably starting at an estimated start position, and determines content underneath the shape function, by employing for example a (small) number of sinusoidal components. This information is contained in the transient code  $C_T$  and more detailed information on generating the transient code  $C_T$  is provided in PCT Patent Application No. WO 01/69593.

The transient code  $C_T$  is furnished to the transient synthesizer 112. The synthesized transient signal component is subtracted from the input signal  $x(t)$  in subtractor 16, resulting in a signal  $x_1$ . A gain control mechanism GC (12) is used to produce  $x_2$  from  $x_1$ .

The signal  $x_2$  is furnished to the sinusoidal coder 13 where it is analyzed in a sinusoidal analyzer (SA) 130, which determines the (deterministic) sinusoidal components. It will therefore be seen that while the presence of the transient analyser is desirable, it is not necessary and the invention can be implemented without such an analyser. Alternatively, as mentioned above, the invention can also be implemented with for example an harmonic complex analyser.

In brief, the sinusoidal coder encodes the input signal  $x_2$  as tracks of sinusoidal components linked from one frame segment to the next. Referring now to Figure 3(a), in the same manner as in the prior art, in the preferred embodiment, each segment of the input signal  $x_2$  is transformed into the frequency domain in a Fourier Transform (FT) unit 40. For each segment, the FT unit provides measured amplitudes  $A$ , phases  $\phi$  and frequencies  $\omega$ . As mentioned previously, the range of phases provided by the Fourier Transform is restricted to  $-\pi \leq \phi < \pi$ . A tracking algorithm (TA) unit 42 takes the information for each segment and

by employing a suitable cost function, links sinusoids from one segment to the next, so producing a sequence of measured phases  $\phi(k)$  and frequencies  $\omega(k)$  for each track.

In contrast to the prior art, according to the present invention the sinusoidal codes  $C_s$  ultimately produced by the analyzer 130 include phase information, and frequency is reconstructed from this information in the decoder.

As mentioned above, however, the measured phase is restricted to a modulo  $2\pi$  representation. Therefore, in the preferred embodiment, the analyzer comprises a phase unwrapper (PU) 44 where the modulo  $2\pi$  phase representation is unwrapped to expose the structural inter-frame phase behaviour for a track  $\psi$ . As the frequency in sinusoidal tracks is nearly constant, it will be seen that the unwrapped phase  $\psi$  will typically be a linearly increasing (or decreasing) function and this makes cheap transmission of phase possible. The unwrapped phase  $\psi$  is provided as input to a phase encoder (PE) 46 which provides as output representation levels  $r$  suitable for being transmitted.

Referring now to the operation of the phase unwrapper 44, as mentioned above, actual phase  $\psi$  and actual frequency  $\Omega$  for a track are related by:

$$\psi(t) = \int_{T_0}^t \Omega(\tau) d\tau + \psi(T_0) \quad \text{Equation 1}$$

with  $T_0$  a reference time instant.

A sinusoidal track in frames  $k = K, K+1 \dots K+L-1$  has measured frequencies  $\omega(k)$  (expressed in radians per second) and measured phases  $\phi(k)$  (expressed in radians). The distance between the centre of the frames is given by  $U$  (update rate expressed in seconds). The measured frequencies are supposed to be samples of the assumed underlying continuous-time frequency track  $\Omega$  with  $\omega(k) = \Omega(kU)$  and, similarly, the measured phases are samples of the associated continuous-time phase track  $\psi$  with  $\phi(k) = \psi(kU) \bmod (2\pi)$ . For sinusoidal coding it is assumed that  $\Omega$  is a nearly constant function.

Assuming that the frequencies are nearly constant within a segment Equation 1 can be approximated as follows:

$$\begin{aligned} \psi(kU) &= \int_{(k-1)U}^{kU} \Omega(t) dt + \psi((k-1)U) \\ &\approx \{\omega(k) + \omega(k-1)\}U / 2 + \psi((k-1)U). \end{aligned} \quad \text{Equation 2}$$

It will therefore be seen that knowing the phase and frequency for a given segment and the frequency of the next segment, it is possible to estimate an unwrapped phase value for the next segment, and so on for each segment in a track.

In the preferred embodiment, the phase unwrapper determines an unwrap factor  $m(k)$  at instant  $k$ :

$$\psi(kU) = \phi(k) + m(k)2\pi \quad \text{Equation 3}$$

The unwrap factor  $m(k)$  tells the phase unwrapper the number of cycles which has to be added to obtain the unwrapped phase.

Combining equations 2 and 3, the phase unwrapper determines an incremental unwrap factor  $e$  as follows:

$$2\pi e(k) = 2\pi \{m(k) - m(k-1)\} = \{\omega(k) + \omega(k-1)\}U / 2 - \{\phi(k) - \phi(k-1)\}$$

where  $e$  should be an integer. However, due to measurement and model errors, the incremental unwrap factor will not be an integer exactly, so:

$$e(k) = \text{round}([\{\omega(k) + \omega(k-1)\}U / 2 - \{\phi(k) - \phi(k-1)\}] / (2\pi))$$

assuming that the model and measurement errors are small.

Having the incremental unwrap factor  $e$ , the  $m(k)$  from equation (3) is calculated as the cumulative sum where, without loss of generality, the phase unwrapper starts in the first frame  $K$  with  $m(K)=0$ , and from  $m(k)$  and  $\phi(k)$ , the (unwrapped) phase  $\psi(kU)$  is determined.

In practice, the sampled data  $\psi(kU)$  and  $\Omega(kU)$  are distorted by measurement errors:

$$\phi(k) = \psi(kU) + \varepsilon_1(k),$$

$$\omega(k) = \Omega(kU) + \varepsilon_2(k),$$

where  $\varepsilon_1$  and  $\varepsilon_2$  are the phase and frequency errors, respectively. In order to prevent the

determination of the unwrap factor becoming ambiguous, the measurement data needs to be determined with sufficient accuracy. Thus, in the preferred embodiment, tracking is restricted so that:

$$\delta(k) = e(k) - [\{\omega(k) + \omega(k-1)\}U / 2 - \{\phi(k) - \phi(k-1)\}] / (2\pi) < \delta_0$$

where  $\delta$  is the error in the rounding operation. The error  $\delta$  is mainly determined by the errors in  $\omega$  due to the multiplication with  $U$ . Assume that  $\omega$  is determined from the maxima of the absolute value of the Fourier Transform from a sampled version of the input signal with

sampling frequency  $F_s$  and that the resolution of the Fourier Transform is  $2\pi/L_a$  with  $L_a$  the analysis size. In order to be within the considered bound, we have:

$$\frac{L_a}{U} = \delta_0$$

That means that the analysis size should be few times larger than the update size in order for unwrapping to be accurate, e.g., setting  $\delta_0 = 1/4$ , the analysis size should be four times the update size (neglecting the errors  $\varepsilon_1$  in the phase measurement).

The second precaution which can be taken to avoid decision errors in the round operation is to defining tracks appropriately. In the tracking unit 42, sinusoidal tracks are typically defined by considering amplitude and frequency differences. Additionally, it is also possible to account for phase information in the linking criterion. For instance, we can define the phase prediction error  $\varepsilon$  as the difference between the measured value and the predicted value  $\tilde{\phi}$  according to

$$\varepsilon = \{\phi(k) - \tilde{\phi}(k)\} \bmod 2\pi$$

where the predicted value can be taken as

$$\tilde{\phi}(k) = \phi(k-1) + \{\omega(k) - \omega(k-1)\}U / 2$$

Thus, preferably the tracking unit 42 forbids tracks where  $\varepsilon$  is larger than a certain value (e.g.  $\varepsilon > \pi/2$ ), resulting in an unambiguous definition of  $e(k)$ .

Additionally, the encoder may calculate the phases and frequencies such as will be available in the decoder. If the phases or frequencies which will become available in the decoder differ too much from the phases and/or frequencies such as are present in the encoder, it may be decided to interrupt a track, i.e. to signal the end of a track and start a new one using the current frequency and phase and their linked sinusoidal data.

The sampled unwrapped phase  $\psi(kU)$  produced by the phase unwrapper (PU) 44 is provided as input to phase encoder (PE) 46 to produce the set of representation levels  $r$ . Techniques for efficient transmission of a generally monotonically changing characteristic such as the unwrapped phase are known. In the preferred embodiment, Figure 3(b), Adaptive Differential Pulse Code Modulation (ADPCM) is employed. Here, a predictor (PF) 48 is used to estimate the phase of the next track segment and encode the difference only in a quantizer (Q) 50. Since  $\psi$  is expected to be a nearly linear function and for reasons of simplicity, the predictor 48 is chosen as a second-order filter of the form:

$$y(k+1) = 2x(k) - x(k-1)$$



where  $x$  is the input and  $y$  is the output. It will be seen, however, that it is also possible to take other functional relations (including higher-order relations) and to include adaptive (backward or forward) adaptation of the filter coefficients. In the preferred embodiment, a backward adaptive control mechanism (QC) 52 is used for simplicity to control the quantiser

5 50. Forward adaptive control is also possible as well but would require extra bit rate overhead.

As will be seen, initialization of the encoder (and decoder) for a track starts with knowledge of the start phase  $\phi(0)$  and frequency  $\omega(0)$ . These are quantized and transmitted by a separate mechanism. Additionally, the initial quantization step used in the

10 quantization controller 52 of the encoder and the corresponding controller 62 in the decoder, Figure 5(b), is either transmitted or set to a certain value in both encoder and decoder. Finally, the end of a track can either be signalled in a separate side stream or as a unique symbol in the bit stream of the phases.

From the sinusoidal code  $C_S$  generated with the sinusoidal coder, the

15 sinusoidal signal component is reconstructed by a sinusoidal synthesizer (SS) 131 in the same manner as will be described for the sinusoidal synthesizer (SS) 32 of the decoder. This signal is subtracted in subtractor 17 from the input  $x_2$  to the sinusoidal coder 13, resulting in a remaining signal  $x_3$ . The residual signal  $x_3$  produced by the sinusoidal coder 13 is passed to the noise analyzer 14 of the preferred embodiment which produces a noise code  $C_N$

20 representative of this noise, as described in, for example, PCT patent application No. PCT/EP00/04599.

Finally, in a multiplexer 15, an audio stream AS is constituted which includes the codes  $C_T$ ,  $C_S$  and  $C_N$ . The audio stream AS is furnished to e.g. a data bus, an antenna system, a storage medium etc.

25 Fig. 4 shows an audio player 3 suitable for decoding an audio stream AS', e.g. generated by an encoder 1 of Fig. 1, obtained from a data bus, antenna system, storage medium etc. The audio stream AS' is de-multiplexed in a de-multiplexer 30 to obtain the codes  $C_T$ ,  $C_S$  and  $C_N$ . These codes are furnished to a transient synthesizer 31, a sinusoidal synthesizer 32 and a noise synthesizer 33 respectively. From the transient code  $C_T$ , the

30 transient signal components are calculated in the transient synthesizer 31. In case the transient code indicates a shape function, the shape is calculated based on the received parameters. Further, the shape content is calculated based on the frequencies and amplitudes of the sinusoidal components. If the transient code  $C_T$  indicates a step, then no transient is calculated. The total transient signal  $y_T$  is a sum of all transients.

The sinusoidal code  $C_S$  including the information encoded by the analyser 130 is used by the sinusoidal synthesizer 32 to generate signal  $y_S$ . Referring now to Figures 5(a) and (b), the sinusoidal synthesizer 32 comprises a phase decoder (PD) 56 compatible with the phase encoder 46. Here, a dequantiser (DQ) 60 in conjunction with a second-order prediction filter (PF) 64 produces (an estimate of) the unwrapped phase  $\hat{\psi}$  from: the representation levels  $r$ ; initial information  $\hat{\phi}(0), \hat{\omega}(0)$  provided to the prediction filter (PF) 64 and the initial quantization step for the quantization controller (QC) 62.

As illustrated in Figure 2(b), the frequency can be recovered from the unwrapped phase  $\hat{\psi}$  by differentiation. Assuming that the phase error at the decoder is approximately white and since differentiation amplifies the high frequencies, the differentiation can be combined with a low-pass filter to reduce the noise and, thus, to obtain an accurate estimate of the frequency at the decoder.

In the preferred embodiment, a filtering unit (FR) 58 approximates the differentiation which is necessary to obtain the frequency  $\hat{\omega}$  from the unwrapped phase by procedures as forward, backward or central differences. This enables the decoder to produce as output the phases  $\hat{\psi}$  and frequencies  $\hat{\omega}$  usable in a conventional manner to synthesize the sinusoidal component of the encoded signal.

At the same time, as the sinusoidal components of the signal are being synthesized, the noise code  $C_N$  is fed to a noise synthesizer NS 33, which is mainly a filter, having a frequency response approximating the spectrum of the noise. The NS 33 generates reconstructed noise  $y_N$  by filtering a white noise signal with the noise code  $C_N$ . The total signal  $y(t)$  comprises the sum of the transient signal  $y_T$  and the product of any amplitude decompression ( $g$ ) and the sum of the sinusoidal signal  $y_S$  and the noise signal  $y_N$ . The audio player comprises two adders 36 and 37 to sum respective signals. The total signal is furnished to an output unit 35, which is e.g. a speaker.

Fig. 6 shows an audio system according to the invention comprising an audio coder 1 as shown in Fig. 1 and an audio player 3 as shown in Fig. 4. Such a system offers playing and recording features. The audio stream AS is furnished from the audio coder to the audio player over a communication channel 2, which may be a wireless connection, a data bus or a storage medium. In case the communication channel 2 is a storage medium, the storage medium may be fixed in the system or may also be a removable disc, memory stick etc. The communication channel 2 may be part of the audio system, but will however often be outside the audio system.

## CLAIMS:

1. A method of encoding an audio signal, the method comprising the steps of:  
providing a respective set of sampled signal values for each of a plurality of sequential segments;  
analysing the sampled signal values to determine one or more sinusoidal  
5 components for each of the plurality of sequential segments;  
linking sinusoidal components across a plurality of sequential segments to provide sinusoidal tracks;  
for each sinusoidal track, determining a phase comprising a generally monotonically changing value; and  
10 generating an encoded audio stream including sinusoidal codes representing said phase.
2. A method according to claim 1 wherein the phase value of each linked segment is determined as a function of: the integral of the frequency for the previous segment  
15 and the frequency of said linked segment; and the phase of the previous segment.
3. A method according to claim 1 wherein said sinusoidal components include: a frequency value; and a phase value in the range  $\{-\pi, \pi\}$ .
- 20 4. A method according to claim 1 wherein generating step comprises:  
predicting a value of phase for a segment as a function of phase for at least the previous segment; and  
quantizing said sinusoidal codes as a function of said predicted value for said phase and the measured phase for said segment.  
25
5. A method according to claim 4 wherein said sinusoidal codes for a track include an initial phase and frequency, and wherein said predicting step employs said initial frequency and phase to provide a first prediction.

6. A method according to claim 4 wherein said generating step comprises:  
controlling said quantizing step as a function of said quantized sinusoidal codes.

7. A method according to claim 6 wherein said sinusoidal codes for each track  
5 include an initial quantization step.

8. A method as claimed in claim 1 wherein said sinusoidal codes include an  
indicator of the end of a track.

10 9. A method according to claim 1 further comprising:  
synthesizing said sinusoidal components using said sinusoidal codes;  
subtracting said synthesized signal values from said sampled signal values to  
provide a set of values representing a remainder component of said audio signal;  
modelling the remainder component of the audio signal by determining  
15 parameters, approximating the remainder component; and  
including said parameters in said audio stream.

10. A method according to claim 1 wherein said sampled signal values represent  
an audio signal from which transient components have been removed.

20 11. Method of decoding an audio stream, the method comprising the steps of:  
reading an encoded audio stream including sinusoidal codes representing a  
phase for each track of linked sinusoidal components,  
for each track, generating a generally monotonically changing value from said  
25 codes representing said phase;  
filtering said generated value to provide an estimate of frequency for a track;  
and  
employing said generated values and said frequency estimates to synthesize  
said sinusoidal components of said audio signal.

30 12. Audio coder arranged to process a respective set of sampled signal values for  
each of a plurality of sequential segments of an audio signal, said coder comprising:  
an analyser for analysing the sampled signal values to determine one or more  
sinusoidal components for each of the plurality of sequential segments;

a linker for linking sinusoidal components across a plurality of sequential segments to provide sinusoidal tracks;

a phase unwrapper for determining, for each sinusoidal track, a phase comprising a generally monotonically changing value; and

5 a phase encoder for providing an encoded audio stream including sinusoidal codes representing said phase.

13. Audio player comprising:

means for reading an encoded audio stream including sinusoidal codes  
10 representing a phase for each track of linked sinusoidal components,

a phase unwrapper for determining, for each track, a generally monotonically changing value from said codes representing said phase;

a filter for filtering said generated value to provide an estimate of frequency for a track; and

15 a synthesizer arranged to employ said generated values and said frequency estimates to synthesize said sinusoidal components of said audio signal.

14. Audio system comprising an audio coder as claimed in claim 12 and an audio player as claimed in claim 13.

20

15. Audio stream comprising sinusoidal codes representing tracks of linked sinusoidal components of an audio signal, said codes representing a generally monotonically changing value corresponding to a phase for each track of linked sinusoidal components.

25 16. Storage medium on which an audio stream as claimed in claim 15 has been stored.

1/4

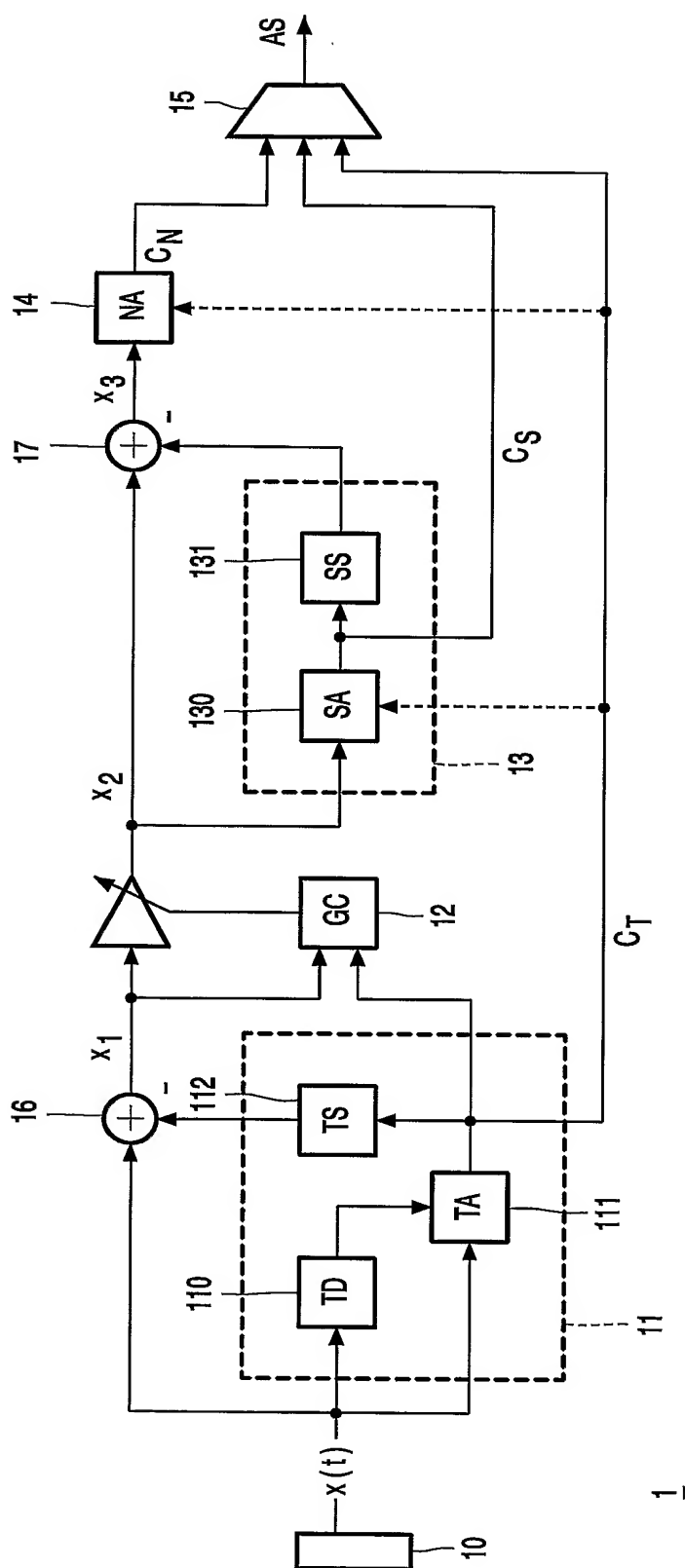


FIG.1

2/4

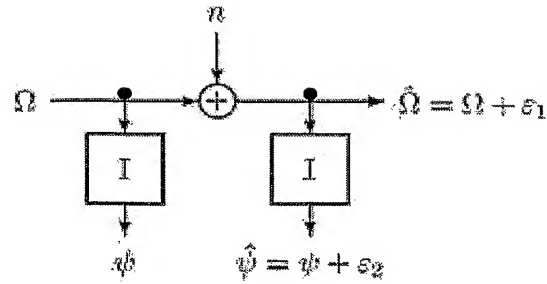


FIG.2a

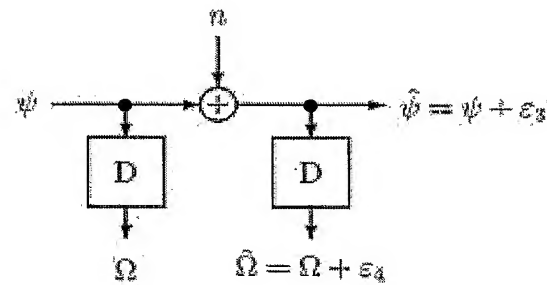


FIG.2b

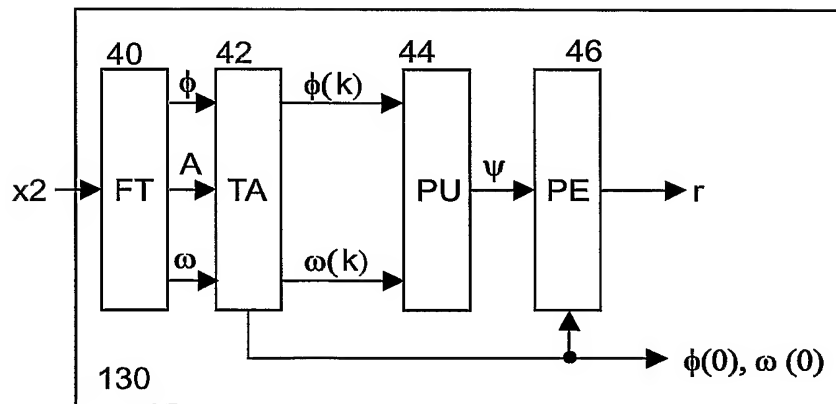


FIG.3a

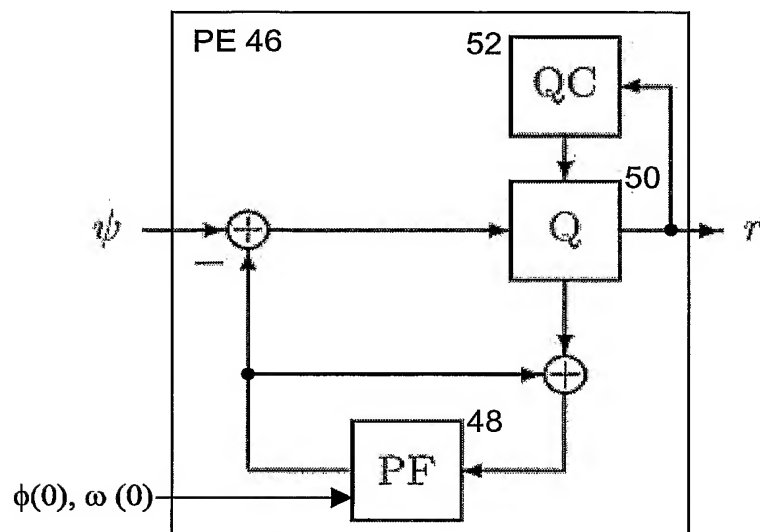


FIG.3b

3/4

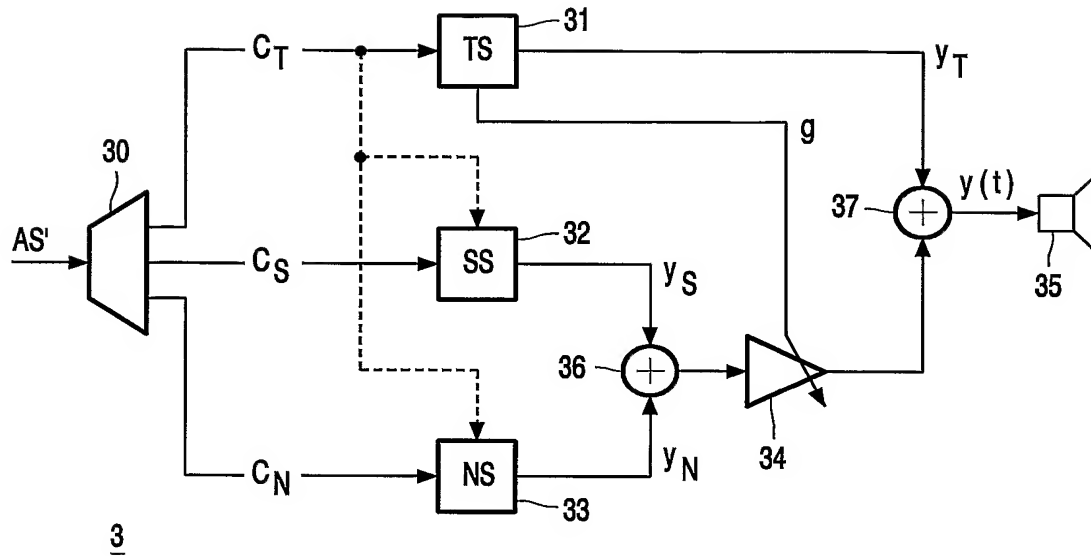


FIG. 4

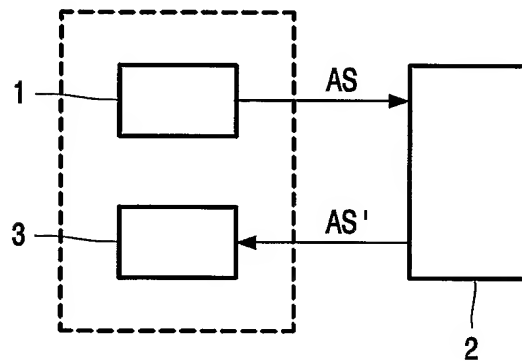


FIG. 6



4/4

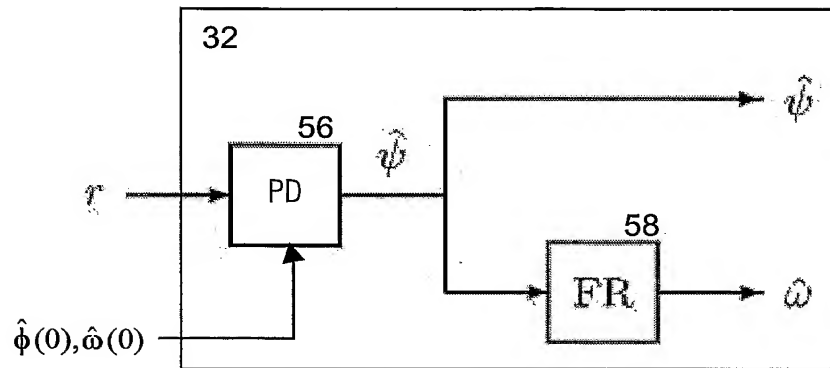


FIG.5a

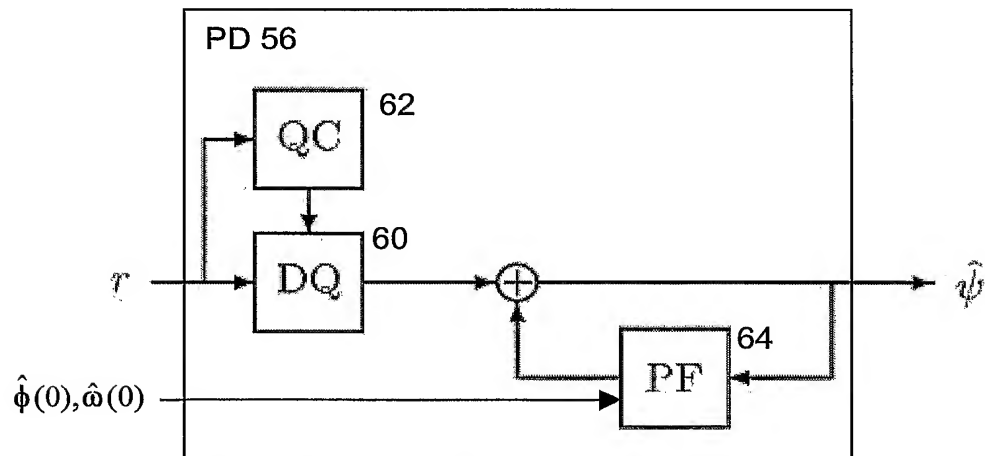


FIG.5b

## INTERNATIONAL SEARCH REPORT

Internati plication No

PCT/IB 03/05019

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 G10L19/08

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)  
IPC 7 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ, INSPEC, COMPENDEX, IBM-TDB

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	AHMADI S ET AL: "Minimum-variance phase prediction and frame interpolation algorithms for low bit rate sinusoidal speech coding" ISCAS 2000 IEEE INTERNATIONAL SYMPOSIUM ON CIRCUITS AND SYSTEMS, vol. 3, 28 May 2000 (2000-05-28), pages 730-733, XP010502635 page 730, left-hand column, line 1 -page 731, right-hand column, line 22	1-8, 12, 15, 16
Y	---	9, 10
	--- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

## ° Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

26 March 2004

Date of mailing of the international search report

02/04/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Zimmermann, E

## INTERNATIONAL SEARCH REPORT

Internati plication No  
PCT/IB 03/05019

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	<p>PUERNHAGEN H: "Advances in parametric audio coding" APPLICATIONS OF SIGNAL PROCESSING TO AUDIO AND ACOUSTICS, 1999 IEEE WORKSHOP ON NEW PALTZ, NY, USA 17-20 OCT. 1999, PISCATAWAY, NJ, USA, IEEE, US, 17 October 1999 (1999-10-17), pages 31-34, XP010365061 ISBN: 0-7803-5612-8 abstract; figure 2 paragraph '03.1!</p> <p>----</p>	9,10
P,X	<p>DEN BRINKER A C, GERRITS A J, SLUIJTER R J: "Phase transmission in a sinusoidal audio and speech coder" 115TH AES CONVENTION, AUDIO ENGINEERING SOCIETY, 10 - 13 October 2003, XP009028272 New York the whole document</p> <p>----</p>	1-16
A	<p>EP 0 337 636 A (AMERICAN TELEPHONE &amp; TELEGRAPH) 18 October 1989 (1989-10-18) abstract page 3, line 7 - line 21</p> <p>----</p>	11,13,14
A	<p>MCAULAY R J ET AL: "SPEECH ANALYSIS/SYNTHESIS BASED ON A SINUSOIDAL REPRESENTATION" IEEE TRANSACTIONS ON ACOUSTICS, SPEECH AND SIGNAL PROCESSING, IEEE INC. NEW YORK, US, vol. ASSP-34, no. 4, August 1986 (1986-08), pages 744-754, XP001002928 ISSN: 0096-3518 abstract; figure 7 page 750, left-hand column, line 14 -right-hand column, line 13</p> <p>----</p>	1,11
A	<p>EDLER B ET AL: "ASAC - ANALYSIS/SYNTHESIS AUDIO CODEC FOR VERY LOW BIT RATES" PREPRINTS OF PAPERS PRESENTED AT THE AES CONVENTION, XX, XX, 11 May 1996 (1996-05-11), pages 1-15, XP001062332 abstract paragraph '0003!</p> <p>----</p>	1,11
A	<p>WO 02 056298 A (KONINKL PHILIPS ELECTRONICS NV) 18 July 2002 (2002-07-18) abstract</p> <p>-----</p>	1

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IB 03/05019

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
EP 0337636	A	18-10-1989	US	5179626 A		12-01-1993
			CA	1336456 C		25-07-1995
			DE	68916831 D1		25-08-1994
			EP	0337636 A2		18-10-1989
			JP	2203398 A		13-08-1990
<hr/>						
WO 02056298	A	18-07-2002	CN	1418362 T		14-05-2003
			EP	1356456 A1		29-10-2003
			WO	02056298 A1		18-07-2002
			US	2002133358 A1		19-09-2002
<hr/>						